SCREENING METHOD FOR GENE VARIATION

[0001]

BACKGROUND OF INVENTION

Field of the Invention

The present invention relates to a variation detection method and a system for it useful for screening of a gene variation etc.

[0002]

Related Background Art

One of the techniques for sequencing nucleic acid etc. or for detecting the sequence is to utilize a DNA array. USP 5,445,934 discloses a DNA array where 100,000 or more oligonucleotide probes are bonded in 1 inch square. Such a DNA array has an advantage that many characteristics can be examined at the same time with a very small sample amount. When a fluorescence-labeled sample is poured onto such a DNA chip, DNA fragments in the sample bind to probes having a complementary sequence fixed on the DNA chip, and only that part can be discriminated by fluorescence to elucidate the sequence of the DNA fragment in the DNA sample.

[0003]

Sequencing By Hybridization (SBH) is a method for examining the base sequence utilizing such a DNA array and the details are described in USP 5,202,231. In the SBH method, all possible sequences of an

oligonucleotide of a certain length are arranged on the substrate, then fully matched hybrids formed by hybridization reaction between probes and the sample DNA are detected. If a set of fully matched hybrids is obtained, the set will give an assembly of overlapping sequences with one base shift being a part of one certain sequence, which sequence is extracted for calling.

[0004]

Not only the SBH method, when complementariness between an oligonucleotide and a sample DNA is examined, it is very difficult to call whether a hybrid was formed or not using one probe for one test item, since the stability of a hybrid differs sequence to sequence, and there is no perfect signal for calling the full complementariness. Science vol. 274 p.610-614, 1996 discloses a method for calling comparing the signal intensity of a perfect match hybrid and the weaker intensities of one-base mismatch hybrids. In this method, 15-mer oligonucleotide probes differing from each other only by one mismatching base at the center of the sequence are prepared, and the fluorescence intensities of the hybrids of the probes are compared, and when the intensity of the full matched hybrid is stronger than that of other hybrids by a predetermined rate, it is called positive.

[0005]

Further, USP 5,733,729 discloses a method using a computer to know a base sequence of a sample from a comparison of fluorescence intensities of obtained hybrids for more accurate calling.

[0006]

However, actual binding strength of a hybrid depends on the GC content etc., and difference of the fluorescence intensity between a full match hybrid and a one-base mismatch hybrid also varies in a considerable range depending on the sequence. Thus, a method for calling whether a sequence is fully complementary to a probe or not, using a 15 mer oligonucleotide probe to compare it with other three probes having one mismatched base at the center thereof, can provide more accuracy if each stability is evaluated theoretically or empirically before comparison.

[0007]

In addition, accurate calling requires precise quantification of signals, and therefore, precision apparatuses such as a confocal laser microscope. Furthermore, in order to measure the fluorescence intensity of a hybrid of every probe and to determine the gene sequence by analyzing the data, a large-scale computer apparatus as well as a detection apparatus for reading the arrays are further required. Therefore this is a big obstacle for ready use of the DNA array.

[0008]

On the other hand, gene diagnosis using such a DNA array may be used in group medical examination, individual gene examination or gene-polymorphism study. In such a case, however, the above described precise measurement and analysis are not always required, where a large amount of samples are rapidly treated at a low cost in order to find out variated samples concerning a specific item from a large number of normal samples. Further, the precision apparatus and analysis as described above will cost much. Accordingly, a concept that screening of the presence or absence of a variation is first performed, and then, detailed examinations are carried out about the samples suspected of variation by the screening will save time and cost.

[0009]

SUMMARY OF THE INVENTION

One object of the present invention is to provide a method suitable for mass screening so as to determine rapidly the presence or absence of a gene variation without need of an expensive apparatus and a complex analysis.

[0010]

The present invention provides a DNA array in which a group of probes which will give strong signals forming hybrids with a normal gene sequence, and a

group of probes having sequences expected to form hybrids with gene variants are separately arranged, on the premise that the base sequence of a normal gene and those of variants have already been established. Furthermore, the above described object is achieved by providing a detection method using such an array.

[0011]

According to one aspect of the present invention, there is provided a method for screening of the presence or absence of variation in a region of a nucleic acid comprising the steps of:

(a) preparing a test nucleic acid corresponding to the region;

(b) preparing a probe having a base sequence fully complementary to a normal sequence of the region, and a plurality of probes each having at least one base not complementary to the normal sequence;

(c) fixing the probes in separate regions on a surface of a substrate to prepare a DNA array substrate;

(d) reacting the test nucleic acid with the probes on the DNA array substrate;

(e) measuring signals in each region totaly where the signals are originated from respective hybrids formed between the test nucleic acid and one of the probes; and

(f) determining the presence or absence of

SUB C2

SUB C3 CONT.

mutation in the test nucleic acid comparing with a histogram pattern of signals of all regions obtained using a normal sample without variation.

[0012]

According to another aspect of the invention, there is provided a DNA array substrate for screening a variation in a region of a nucleic acid, wherein

a full match probe fully complementary to a normal sequence of the region, and a plurality of mismatch probes having at least one base mismatch to the sequence are arranged on the substrate; and

the probes are arranged to form at least two separate regions selected from:

a first region containing at least one probe which provides a signal of a certain intensity on reaction with a nucleic acid having the normal sequence,

a second region containing at least one probe which provides a weaker signal than the probe of the first region on reaction with a nucleic acid having normal sequence, and

the third region containing at least one probe which provides no signal on reaction with a nucleic acid having normal sequence.

[0013]

According to still another aspect of the present invention, there is provided a system for detecting variation comprising a DNA array substrate as described

above and a signal measuring apparatus which measures signals from separate regions of the DNA array substrate.

[0014]

The present invention can provide a method suitable for mass screening, so as to rapidly determine only the presence or absence of a gene variant, without need of an expensive apparatus and complex analysis.

[0015]

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A is a schematic example of arrangement of separate regions on a DNA array substrate, and FIG. 1B shows a histogram of fluorescence intensities (pattern) corresponding to separate regions of the above arrangement, measured using a line sensor:

FIG. 2 shows an arrangement of 64 types of DNA probes;

FIG. 3 shows a pattern of the fluorescence intensities obtained by hybridization;

FIG. 4 shows normal fluorescence intensities obtained by hybridization where Gr. NO. denotes group number;

FIG. 5 is a distribution pattern of fluorescence intensities corresponding to the separate regions of the arrangement, measured using a line sensor;

FIG. 6 is a distribution pattern of fluorescence intensities corresponding to the separate regions of

the arrangement, measured using a line sensor;

FIG. 7 is a distribution pattern of fluorescence intensities corresponding to the separate regions of the arrangement, measured using a line sensor;

FIG. 8 is a distribution pattern of fluorescence intensities corresponding to the separate regions of the arrangement, measured using a line sensor; and

FIG. 9 shows normal fluorescence intensities obtained by hybridization where Gr. NO. denotes group number.

[0016]

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a screening method for gene variants using a DNA array in which a group of probes which will give strong signals forming hybrids with a normal gene sequence, and a group of probes having sequences expected to form hybrids with gene variants are separately arranged, on the premise that the base sequence of a normal gene and those of variants have already been established. Furthermore, the above described object is achieved by providing a detection method using such an array.

[0017]

Here, the present invention will be described in detail with examples where signals from the separate regions are fluorescence. However, signals in the present invention are not limited to fluorescence but

may be other light signals or electric signals.

Binding strength between single-stranded nucleic acids to form a hybrid is controlled by various factors, and when a probe having a length of about 12 mer to 25 mer is used, it is practically difficult to perfectly exclude hybrids having one-base mismatches.

[0018]

When the signal of the hybrid to be detected is a light such as fluorescence, the following phenomena are observed. Fluorescence stability of a hybrid having two mismatches (two-base mismatch hybrids) is much lower than that of the one-base mismatched hybrid, regardless of positions, continuity or discontinuity of the mismatched bases. On the other hand, signal from a three-base mismatch hybrid is hardly observed. However, one-base mismatch hybrids may have more than 50% signal intensity of a full match hybrid. Thus, when a sample is a nucleic acid of normal sequence, strong fluorescence is observed in a region where a full match probe and probes having one mismatching base have been arranged, while the fluorescence intensity in a region where hybrids of low stability are formed is almost zero. On the other hand, when a sample is a nucleic acid of a variant sequence, the probe fully complementary to the normal sample makes a mismatch hybrids with the sample. Thus the fluorescence is weaker at a region than that in case of the normal

sample, at the same time, fluorescence from a full match hybrid and one-base mismatch hybrids with the sample appears in another region where low signals are expected in case of the normal nucleic acid. Accordingly, by comparing fluorescence histogram of the regions, it is possible to distinguish between the normal nucleic acid and the variant nucleic acid.

[0019]

In the present invention, a DNA array substrate where probes are arranged in separate regions according to the fluorescence intensities of their hybrids with a normal nucleic acid is used for more accurate calling. First, hybridization reaction of the normal nucleic acid with each probe is performed, and based on the fluorescence intensities of the hybrids obtained, separate regions each containing probes corresponding to strong fluorescence, no fluorescence, and moderate (weak) fluorescence are located at predetermined positions on the substrate.

[0020]

When performing a test on multiple items simultaneously, the substrate should be divided into areas for respective items. In each area, probes having full match and one-base mismatch sequences to the normal gene sequence are arranged in a region, which are expected to give high signals, and the other probes having more than two-base mismatch are arranged

in a separate region depending on the items. Thus we can discriminate normal test samples from variated ones.

[0021]

The arrangement can be determined according to the type of the sensor to be used. For example, when a line sensor is used, each region is arranged from the left to right in the substrate in order of the strength of fluorescence obtainable by hybridization with the normal nucleic acid. Thus, the fluorescence intensity will be maximum in the left, then gradually decrease and come to zero in several regions in the right of the substrate. When an area sensor is used, it is necessary to evaluate the fluorescence quantity of at least two separate regions containing a group of probes which will provide the maximum fluorescence and a group of probes which will not provide any fluorescence respectively.

[0022]

This will be described more specifically.

[0023]

In the present invention, for example, whether a gene is normal or not can be called by providing a region containing a probe that forms a hybrid with a normal nucleic acid and the other region containing probes that form hybrids with variant genes separately on a DNA array, and taking a ratio of the signal from a

hybrid corresponding to the normal nucleic acid and to the signals from hybrids corresponding to variant genes.

[0024]

However, because one-base mismatched hybrids sometimes have strong signals, it is difficult to detect variation in the test sample when a DNA array where only probes being full match or one-base mismatch to a normal sequence are arranged is used. It is important that probes having two-base mismatch to the normal nucleic acid is present in the DNA array, which might be one-base mismatch probes to the test sample to form hybrids having strong signals.

[0025]

Therefore, in the present invention, in order to judge more accurately, a preferable method is as follows. Probes having full match and one-base mismatch sequences to the normal nucleic acid sequence, which most samples in mass screening have, are arranged in a specified region on a DNA array substrate. In addition, probes having two- or three-base mismatch are arranged in a region different from the above region for high stability hybrids. Then, a hybridization reaction is performed with the normal nucleic acid or a sample, using the DNA array substrate of such an arrangement. Then the total of signals for each separate region is measured, and a pattern obtained

with a sample nucleic acid is compared that with the normal nucleic acid are compared to determine the presence or absence of variation.

[0026]

For example, when 64 probes (4 × 4 × 4 = 64) where different bases are arranged at three positions in 18 base length (Table 1, 5'-terminus is on the left) are used supposing that variation occurs at these three points, for any sample there should be present a probe fully complementary to the sample, nine one-base mismatch probes, 27 two-base mismatch probes and 27 three-base mismatch probes. Furthermore, it is relatively easy to set conditions of the hybridization reaction such that fluorescence is observed with the full match and one-base mismatch hybrids but not with three-base mismatch hybrids. Some two-base mismatch probes form hybrids and others not, depending to their sequences.

[0027]

[Table 1]

SEQ ID NO:	Sequence	SEQ ID NO:	Sequence
1	GATGGGACTCAAGTTCAT	33	GATGGGCCTCAAGTTCAT
2	GATGGGACTCAGGTTCAT	34	GATGGGCCTCAGGTTCAT
3	GATGGGACTCACGTTCAT	35	GATGGGCCTCACGTTCAT
4	GATGGGACTCATGTTCAT	36	GATGGGCCTCATGTTCAT
5	GATGGGACTCGAGTTCAT	37	GATGGGCCTCGAGTTCAT
6	GATGGGACTCGGGTTCAT	38	GATGGGCCTCGGGTTCAT
7	GATGGGACTCGCGTTCAT	39	GATGGGCCTCGCGTTCAT
8	GATGGGACTCGTGTTCAT	40	GATGGGCCTCGTGTTCAT
9	GATGGGACTCCAGTTCAT	41	GATGGGCCTCCAGTTCAT
10	GATGGGACTCCGGTTCAT	42	GATGGGCCTCCGGTTCAT
11	GATGGGACTCCCGTTCAT	43	GATGGGCCTCCCGTTCAT
12	GATGGGACTCCTGTTCAT	44	GATGGGCCTCCTGTTCAT
13	GATGGGACTCTAGTTCAT	45	GATGGGCCTCTAGTTCAT
14	GATGGGACTCTGGTTCAT	46	GATGGGCCTCTGGTTCAT
15	GATGGGACTCTCGTTCAT	47	GATGGGCCTCTCGTTCAT
16	GATGGGACTCTTGTTCAT	48	GATGGGCCTCTTGTTCAT
17	GATGGGGCTCAAGTTCAT	49	GATGGGTCTCAAGTTCAT
18	GATGGGGCTCAGGTTCAT	50	GATGGGTTCTAGGTTCAT
19	GATGGGGCTCACGTTCAT	51	GATGGGTCTCACGTTCAT
20	GATGGGGCTCATGTTCAT	52	GATGGGTCTCATGTTCAT
21	GATGGGGCTCGAGTTCAT	53	GATGGGTCTCGAGTTCAT
22	GATGGGGCTCGGGTTCAT	54	GATGGGTCTCGGGTTCAT
23	GATGGGGCTCGCGTTCAT	55	GATGGGTCTCGCGTTCAT
24	GATGGGGCTCGTGTTCAT	56	GATGGGTCTCGTGTTCAT
25	GATGGGGCTCCAGTTCAT	57	GATGGGTCTCCAGTTCAT
26	GATGGGGCTCCGGTTCAT	58	GATGGGTCTCCGGTTCAT
27	GATGGGGCTCCCGTTCAT	59	GATGGGTCTCCCGTTCAT
28	GATGGGGCTCCTGTTCAT	60	GATGGGTCTCCTGTTCAT
29	GATGGGGCTCTAGTTCAT	61	GATGGGTCTCTAGTTCAT
30	GATGGGGCTCTGGTTCAT	62	GATGGGTCTCTGGTTCAT
31	GATGGGGCTCTCGTTCAT	63	GATGGGTCTCTCGTTCAT
32	GATGGGGCTCTTGTTCAT	64	GATGGGTCTCTTGTTCAT

[0028]

When these 64 probes are grouped into every eight probes in order of the fluorescence intensity obtained by hybridization with the normal nucleic acid (hereinafter "fluorescence intensity of a probe(s)" means expected inensity of a hybrid of the probe with a nucleic acid of normal sequence, if not otherwise stated), the total fluorescence intensity of the first group should be extremely high and the total fluorescence quantity of the sixth, seventh and eighth groups should be zero. Such classification by the fluorescence intensity may be performed empirically or theoretically through calculation.

SUB C3

[0029]

Then, the hybridization reaction is performed in an actual system and the total quantity of fluorescence is determined for each classified group. Particularly, the fluorescence quantities of the sixth, seventh and eighth groups are important. Normally, fluorescence is not expected for these groups. When fluorescence is detected unexpectedly for these groups, it is understood that the sample is not normal but having variation at a part of the sequence. When the fluorescence is not detected for these groups, most of the sequences are normal.

[0030]

For more accuracy, however, it is necessary to

compare the fluorescence quantity of the first group with a normal case. When the fluorescence quantity is significantly lower than that expected for normal sequence, base variation is suspected. Further the histogram of the fluorescence intensity in the regions having medium fluorescence intensity, i.e. the second, third and fourth groups, should be compared with that of the normal one.

[0031]

For more effective detection of variation, one may consider a method wherein the probes are arranged in order of signal intensity, that is, lined in the order of sequences of full match, one-base mismatch, two-base mismatch, and three-base mismatch to the normal sequence from one end, so that signal intensity may be zero in the region opposite to the region where the full match probe is placed. In this case, the total distribution pattern of the fluorescence intensity is examined. For example, when using an arrangement of the separate regions shown in FIG. 1A (one column corresponds to one separate region), a monotonously decreasing pattern as shown in FIG. 1B is obtained for the normal case while patterns having peaks in the intermediate regions will be obtained with variants.

[0032]

The above arrangement method is similarly used for the case where the probes for multi-item testing are

arranged on the same substrate. For each item, first, the full matched probe (to the normal sequence), then one-base mismatched sequences, two-base mismatched sequences and so on are placed in order of the strength of fluorescence intensity expected.

[0033]

Such concept is universally applicable to any number of variation, not limited to the above method where the variation for only three bases is tested.

[0034]

In addition, while we explained the cases where the signals can be obtained when the hybrids are formed, the method may be set such that signals are not obtained when the hybrids are formed, and obtained when the hybrids are not formed, depending on the signal generating system.

[0035]

The sample nucleic acid can be prepared by extracting from the gene to be tested according to the necessity. The control normal nucleic acid can be synthesized on the basis of the known sequence.

[0036]

The length of the probe fixed to the substrate is not limited so long as it is suitable for detection, for example, preferably 8 mer to 30 mer, more preferably 12 mer to 25 mer.

[0037]

Probes may be fixed to the substrate by various methods, but droplet application by the ink jet method is preferably used in order to arrange spots of fixed probes with high efficiency, high density and high speed.

Each spot in the separate region is preferably 70 to 100 μm in diameter and spaced not to connect each other.

[0038]

The spot number in each region is set so that the spots expected to have high fluorescence intensity can be measured together and the spots expected to have no fluorescence can be measured together, considering the constitution of the sensor.

[0039]

A system for detecting variation of the present invention comprises a DNA array substrate wherein plural separate regions are arranged in a prescribed arrangement, and a sensor for measuring the signals in the separate regions subjected to measurement. In addition, for calling the presence or absence of variation using the DNA array substrate where the plural separate regions are provided, the signals from all separate regions may be detected and the results are used for calling the presence or absence of variation or, as a simpler method, signals may be compared between certain regions selected so as to

detect the presence or absence of variation.

[0040]

Computer analysis is conveniently performed connecting a computer system to the detection system. For many variants of a gene sequence, histogram patterns of signal intensity are prepared and stored in the computer, which helps the determination of variant easily and correctly.

[0041]

As the sensor for signal detection, different types of photodiodes (e.g. Hamamatsu Photonics K.K.-made) are used. For example, some of the divided type silicon photodiode arrays can be used as both line sensor and area sensor. Particularly, those having a photo-receiving face of 1 to 2 mm × (2 to 3) mm are suitably used.

[0042]

Examples

The present invention will be described in more detail referring to Examples below. Herein, "%" means "weight %".

Example 1

Preparation of DNA array for detection of variant gene for line sensor

1) Preparation of DNA array linked with 64 types of probes

(1) Probe Design

It is well known that in the base sequence CGGAGG corresponding to the AA248 and AA249 of the tumor suppressor gene p53, frequently observed variation is those the first C to T, the second A to G for AA248, and the third G to T for AA249. Accordingly, aiming at these three positions, 64 types of probes were designed.

[0043]

Sub B1

That is, the designed nucleic acid are 18-mer nucleic acids harboring variegated above mentioned six bases sandwiched between the common sequences, represented by 5'ATGAACNNGAGNCCCATC3' where N corresponds to any of 4 bases, A, G, C and T. Actual probes to detect the above sequence should be have a complementary sequence of 5'GATGGGNCTCNNGTTCAT3'.

[0044]

FIG. 2 shows an arrangement of 64 types of DNA probes on a substrate. A sequence 5' ATGAACCGGAGGCCCATC3' corresponding to the normal gene is expected to form a hybrid with DNA of probe 42 of 5'GATGGGCCTCCGGTTCAT3' positioned at the third from the right and the third from the top.

(2) Preparation of substrate introduced with maleimide group

Substrate Cleaning

A glass plate of 1 inch square was placed in a rack and soaked in an ultrasonic cleaning detergent

overnight. Then, after 20 min of ultrasonic cleaning, the detergent was removed by washing with water. After rinsing the plate with distilled water, ultrasonic treatment was repeated in a container filled with distilled water, for additional 20 min. Then the plate was soaked in a prewarmed 1N sodium hydroxide solution for 10 min, washed with water and then distilled water.

Surface treatment

Then the plate was soaked in an aqueous solution of 1% silane coupling agent (product of Shin-Etsu Chemical Industry: Trade name KBM 603) at a room temperature for 20 min, thereafter nitrogen gas was blown on the both sides blowing off water to dryness. The silane coupling treatment was completed by baking the plate in an oven at 120°C for 1 hour. Subsequently, 2.7 mg of EMCS (N-(6-maleimidocaproyloxy) succinimide: Dojin Company) was weighed and dissolved in a 1 : 1 solution of DMSO/ethanol (final concentration: 0.3 mg/ml). The glass substrate treated with the silane coupling agent was soaked in this EMCS solution for 2 hours to react the amino group of the silane coupling agent with the succimide group of EMCS. At this stage, the maleimide group of EMCS is transferred to the glass surface. After that, the glass plate was washed with ethanol, and dried with nitrogen gas to be used for a coupling reaction with DNA.

3. Coupling of DNA to the substrate

Synthesis of 64 DNA probes

The 64 types of probe DNAs shown in Table 1 each having an SH group (thiol group) at the 5' terminus were synthesized by a standard method.

Ejection of DNA probes

Each DNA was dissolved in water and diluted with SG Clear (aqueous solution containing 7.5% of glycerin, 7.5% of urea, 7.5% of thiodiglycol and 1% of acetylenol EH), to a final concentration of 8 μM.

[0045]

Then 100 μl of this DNA solution was filled into a nozzle of a BJ printer Head BC 62 (Canon) modified to eject a small amount, and to eject six solutions per head. Two heads were used at a time so that 12 types of DNAs could be ejected at once, and the heads were changed 6 times so that 64 spots of 64 types of DNAs were formed independently on the predetermined positions. Thus obtained was a DNA array in which separate regions were arranged in a predetermined manner. The pitch of spots was 200 μm and the area formed with 8 × 8 spots was about 2 mm × 2 mm.

[0046]

After that, the plate was left standing in a humidified chamber for 30 min for linking reaction of the probe DNA to the substrate.

2. Measurement of fluorescence intensity of 64 hybrids

with normal p53 sequence

(1) Hybridization reaction

- Blocking reaction

After completion of the reaction, the substrate was washed with a 1 M NaCl/50 mM phosphate buffer solution (pH 7.0) to wash out thoroughly the DNA solution on the glass surface. Then, this was soaked in an aqueous solution of 2% bovine serum albumin and allowed to stand for 2 hours to carry out a blocking reaction.

- Preparation of model sample DNA

Rhodamine labeled DNA No. 1 (SEQ ID NO: 65) of the same length as the probes but having the normal sequence of p53 gene was prepared. The sequence is shown below and rhodamine is bonded to the 5' terminus.(Synthesis of model sample of DNA)

The labeled DNA No. 65 (single strand) having the normal sequence of p53 gene (complementary to No. 42) and the same length in the same region as the probe DNA was prepared. The sequence is as shown below where rhodamine (Rho) is bound to the 5'-terminal.

Sub B2 No. 65: 5'-Rho-ATGAACCGGAGGCCCATC-3'

~~-Reaction condition of hybridization~~

Two milliliters of 50 nM DNA solution of a model sample containing 100 mM NaCl was placed into a container containing the DNA array substrate for a hybridization reaction. Initially it was heated at

70° C for 30 min, then the temperature of the incubator was lowered to 40° C and the reaction was continue for 3 hours.

(2) Detection

- Method

The detection was performed by connecting an image analysis processing apparatus, ARGUS (a product of Hamamatsu Photonics) to a fluorescence microscope (a product of Nicon).

(3) Result

Distribution of the fluorescence quantity on the substrate obtained is shown in FIG. 3.

[0047]

Example 2

Preparation of DNA array for detection of variant gene for line sensor

(1) Preparation of DNA array for detection of variation

In Tables 2 and 3, these 64 probes are grouped in every 8 probes in order of intensity based on the above described results. The fluorescence intensity of the first group should be extremely strong, and the total fluorescence quantity of the sixth, seventh and eighth groups should be zero.

[0048]

[Table 2]

Group No.	SEQ ID NO:	Type of Mismatch
1	42	Full mismatch
	58	One-base mismatch
	34	One-base mismatch
	41	One-base mismatch
	46	One-base mismatch
	10	One-base mismatch
	44	One-base mismatch
	43	One-base mismatch
2	26	One-base mismatch
	38	One-base mismatch
	50	Two-base mismatch
	2	Two-base mismatch
	6	Two-base mismatch
	9	Two-base mismatch
	11	Two-base mismatch
	12	Two-base mismatch
3	14	Two-base mismatch
	18	Two-base mismatch
	22	Two-base mismatch
	25	Two-base mismatch
	27	Two-base mismatch
	28	Two-base mismatch
	30	Two-base mismatch
	33	Two-base mismatch
4	35	Two-base mismatch
	36	Two-base mismatch
	37	Two-base mismatch
	39	Two-base mismatch
	40	Two-base mismatch
	45	Two-base mismatch
	47	Two-base mismatch
	48	Two-base mismatch

[0049]

[Table 3]

Group No.	SEQ ID NO:	Type of Mismatch
5	54	Two-base mismatch
	57	Two-base mismatch
	59	Two-base mismatch
	60	Two-base mismatch
	62	Two-base mismatch
	1	Three-base mismatch
	3	Three-base mismatch
	4	Three-base mismatch
6	5	Three-base mismatch
	7	Three-base mismatch
	8	Three-base mismatch
	13	Three-base mismatch
	15	Three-base mismatch
	16	Three-base mismatch
	17	Three-base mismatch
	19	Three-base mismatch

[Table 3 (continued)]

Group No.	SEQ ID NO:	Type of Mismatch
7	20	Three-base mismatch
	21	Three-base mismatch
	23	Three-base mismatch
	24	Three-base mismatch
	29	Three-base mismatch
	31	Three-base mismatch
	32	Three-base mismatch
	49	Three-base mismatch
8	51	Three-base mismatch
	52	Three-base mismatch
	53	Three-base mismatch
	55	Three-base mismatch
	56	Three-base mismatch
	61	Three-base mismatch
	63	Three-base mismatch
	64	Three-base mismatch

[0050]

Then, the surface of the substrate was separated into eight columns to arrange the first group, the second group and so on in order of intensity from the left to the right. Then, the probes were arranged at

the positions of the corresponding probe numbers as shown in FIG. 4 using the ink jet method in the same manner is in Example 1 to prepare the DNA arrays for detecting variation. The lines composed of each group have intervals of 200 μm.

(2) Testing normal gene using DNA array

The hybridization reaction was carried out under the same conditions as in Example 1. Thereafter, the total fluorescence of each group was detected using a line sensor (S 272102: Hamamatsu Photonics K.K.-made).

[0051]

The results are shown in FIGS. 4 and 5. The fluorescence of the first group was the highest and the intensity of the second group decreases below the half of that. Fluorescence was hardly observed in the sixth, seventh and eighth groups.

Example 3

Detection of variant gene using DNA array (1)

1. Synthesis of model variant DNA

The labeled DNA No. 66 having the same length as the probes and a sequence complementary to No. 46 probe that differs by one base from the normal sequence of p53 gene was prepared. The sequence is shown below. Rhodamine was bound to the 5'-terminal. The underlined part is the variant position.

No. 66: 5'-Rho-ATGAACCAGAGGCCCATC-3' (SEQ ID NO:66)

Reaction conditions for hybridization

The hybridization reaction was carried out under the same conditions as in Example 1. The concentration of the model sample DNA was 50 nM.

Detection by line sensor

After the hybridization reaction, the DNA array was evaluated in the same manner as in Example 2 using a line sensor. As shown in FIG. 6, the fluorescence was observed in the sixth, seventh and eighth groups where the fluorescence was not observed with the normal gene. Thus, it was easily judged that the sample gene was not normal. Further, since the fluorescence intensity was high at the first and the fourth groups corresponding to probe Nos. 35 to 48, it is inferred that they have the variation in the upper right quarter of FIG. 2 with a sequence very close to the normal gene.

[0052]

This result shows that the screening method for variant genes of the present invention is extremely effective.

[0053]

Example 4

Detection of variant gene using DNA array (2)

The hybridization reaction was carried out under the same conditions as in Example 3, except that the concentration of the model target gene used for the

hybridization reaction was changed to 5 nM. The result is shown in FIG. 7.

[0054]

Since the results similar to Example 3 were obtained to show that the method of the present invention works not depending on the hybridization reaction conditions.

[0055]

Example 5

Detection of variant gene using DNA array (3)

Synthesis of variant model sample DNA

The labeled DNA No. 67 having the same length as the probes and a sequence complementary to No. 10 probe that differs by one base from the normal sequence of p53 gene was prepared. The sequence is as shown below where Rho represents rhodamine bound to the 5'-terminus. The underlined part is the variant position.

No. 67: 5'-Rho-ATGAACCGGAGTCCCATC-3' (SEQ ID NO: 67)

Hybridization reaction

The hybridization reaction was carried out under the same conditions as in Example 1 using the above variation model sample DNA. The concentration of the sample was 50 nM. The result is shown in FIG. 8.

[0056]

Fluorescence was observed in the sixth and seventh groups which was not observed for the normal gene,

showing that this was not a normal gene. Since the fluorescence in the second group was higher than in the first group which would be the highest for the normal gene, it is presumed that it has a variation included in the probe sequences of the second group (the upper left quarter in FIG. 2).

[0057]

This result shows that the screening method for variant genes of the present invention is feasible not depending on the variation types.

[0058]

Example 6

Preparation of DNA array for area sensor

The probes grouped as shown in Example 2 were arranged by the ink jet method in the separate regions on the substrate as shown in FIG. 9 to prepare the DNA array for area sensor.

[0059]

Then, the hybridization reaction was performed using the variant model sample used in Example 4. As a result, the fluorescence was observed in the sixth, seventh and eighth groups which was not observed for the normal gene, therefore, it was easily judged that this gene was a variant one.

SEQUENCE LISTING

<110>Canon INC.

<120>Screening method for gene variation

<130>CFO 15717

<150>JP 2000-263396

<160>67

<210>1

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>1

gatgggactc aagttcat

<210>2

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>2

gatgggactc aggttcat

<210>3

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>3

gatgggactc acgttcat

<210>4

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>4

gatgggactc atgttcat

<210>5

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>5

gatgggactc gagttcat

<210>6

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>6

a gatgggactc ~~gggtcat~~ gggttcat

<210>7

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>7

gatgggactc gcgttcat

<210>8

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>8

gatgggactc gtgttcat

<210>9

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>9

gatgggactc cagttcat

<210>10

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>10

gatgggactc cggttcat

<210>11

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>11

gatgggactc ccgttcat

<210>12

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>12

gatgggactc ctgttcat

<210>13

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>13

gatgggactc tagttcat

<210>14

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>14

gatgggactc tggttcat

<210>15

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>15

gatgggactc tcgttcat

<210>16

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>16

gatgggactc ttgttcat

<210>17

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>17

gatggggctc aagttcat

<210>18

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>18

gatggggctc aggttcat

<210>19

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>19

gatggggctc acgttcat

<210>20

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>20

gatggggctc atgttcat

<210>21

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>21

gatggggctc gagttcat

<210>22

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>22

gatggggctc gggttcat

<210>23

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>23

gatggggctc gcgttcat

<210>24

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>24

gatggggctc gtgttcat

<210>25

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>25

gatggggctc cagttcat

<210>26

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>26

gatggggctc cggttcat

<210>27

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>27

gatggggctc ccgttcat

<210>28

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>28

gatggggctc ctgttcat

<210>29

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>29

gatggggctc tagttcat

<210>30

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>30

gatggggctc tggttcat

<210>31

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>31

gatggggctc tcgttcat

<210>32

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>32

gatggggctc ttgttcat

<210>33

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>33

gatgggcctc aagttcat

<210>34

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>34

gatgggcctc aggttcat

<210>35

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>35

gatgggcctc acgttcat

<210>36

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>36

gatgggcctc atgttcat

<210>37

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>37

gatgggcctc gagttcat

<210>38

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>38

gatgggcctc gggttcat

<210>39

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>39

gatgggcctc gcgttcat

<210>40

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>40

gatgggcctc gtgttcat

<210>41

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>41

gatgggcctc cagttcat

<210>42

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>42

gatgggcctc cggttcat

<210>43

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>43

gatgggcctc ccgttcat

<210>44

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>44

gatgggcctc ctgttcat

<210>45

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>45

gatgggcctc tagttcat

<210>46

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>46

gatgggcctc tggttcat

<210>47

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>47

gatgggcctc tcgttcat

<210>48

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>48

gatgggcctc ttgttcat

<210>49

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>49

gatgggtctc aagttcat

<210>50

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>50

gatgggtctc aggttcat

<210>51

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>51

gatgggtctc acgttcat

<210>52

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>52

gatgggtctc atgttcat

<210>53

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>53

gatgggtctc gagttcat

<210>54

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>54

gatgggtctc gggttcat

<210>55

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>55

gatgggtctc gcgttcat

<210>56

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>56

gatgggtctc gtgttcat

<210>57

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>57

gatgggtctc cagttcat

<210>58

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>58

gatgggtctc cggttcat

<210>59

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>59

gatgggtctc ccgttcat

<210>60

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>60

gatgggtctc ctgttcat

<210>61

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>61

gatgggtctc tagttcat

<210>62

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>62

gatgggtctc tggttcat

<210>63

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>63

gatgggtctc tcgttcat

<210>64

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>64

gatgggtctc ttgttcat

<210>65

<211>18

<212>DNA

<213>p53 fragment

<220>

<223>Sample oligonucleotide

<400>65

atgaaccgga ggcccatc

<210>66

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>66

atgaaccaga ggcccatc

<210>67

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>67

atgaaccgga gtcccatc